

MLLNVLRICI	IVCLVNDGAG	KHSEGRERTK	TYSLNSRGYF	40
RKERGARRSK	ILLVNTKGLD	EPHIGHGDFG	LVAELFDSTR	80
THTNRKEPDM	NKVKLFSTVA	HGNKSARRKA	YNGSRRNIFS	120
RRSFDKRNTE	VTEKPGAKMF	WNNFLVKMNG	APQNTSHGSK	160
AQEIMKEACK	TLPFTQNIVH	ENCDRMVIQN	NLCFGKCISL	200
HVPNQDRRN	TCSHCLPSKF	TLNHLTLNCT	GSKNVVKVVM	240
MVEECTCEAH	KSNFHQTAQF	NMDTSTTLHH		270

Figure 1. Deduced amino acid sequence of *Xenopus cerberus* protein. SEQ ID NO:1.

0990333.074104

MSRTRKVDL LLLAIPGLAL LLLPNAYCAS CEPVRIPMCK SMPWNMTKMP NHLHHSTQAN 60
AILAIEQFEG LLTTECSQDL LFFLCAMYAP ICTIDFQHEP IKPCKSV CER ARAGCEPILI 120
KYRHTWPESL ACEELPVYDR GVCISPEAIV TVEQGTDSMP DFSMDSNNGN CGSGREHCKC 180
KPMKATQKTY LKNYNYVIR AKVKEVKVVC HDATAIVEVK EILKSSLVNI PKDTVTLTYN 240
SGCLCPQLVA NEEYIIMGYE DKERTRLLLV EGSLAEKWRD RLAKKVKRWD QKLRRPRKSK 300
DPVAPIPNKN SNSRQARS

Figure 3. Deduced amino acid sequence of *Xenopus* frazzled protein. SEQ ID NO:3.

Figure 4. Nucleotide sequence of the full-length frazzled cDNA derived from the *Xenopus* organizer. The sense strand of the DNA on top (5' to 3' direction) and the antisense strand on the bottom line (opposite direction). SEQ ID NO:4.

GAATTCCTTT	TCACACAGGA	CTCCTGGCAG	AGGTGAATGG	TTAGCCCTAT	GGATTTGGTT	60
CTTAAGGGAA	AGTGTGTCCT	GAGGACCGTC	TCCACTTACC	AATCGGGATA	CCTAAACCAA	
TGTTGATTTT	GACACATGAT	TGATTGCTTT	CAGATAGGAT	TGAAGGACTT	GGATTTTTAT	120
ACAACTAAAA	CTGTGTACTA	ACTAACGAAA	GTCTATCCTA	ACTTCCTGAA	CCTAAAAATA	
CTAATTCTGC	ACTTTTAAAT	TATCTGAGTA	ATTGTTTATT	TTGTATTGGA	TGGGACTAAA	180
GATTAAGACG	TGAAAATTTA	ATAGACTCAT	TAACAAGTAA	AACATAACCT	ACCTGATTTT	
GATAAACTTA	ACTCCTTGCT	TTTGACTTGC	CCATAAACTA	TAAGGTGGGG	TGAGTTGTAG	240
CTATTTGAAT	TGAGGAACGA	AAACTGAACG	GGTATTTGAT	ATTCCACCCC	ACTCAACATC	
TTGCTTTTAC	ATGTGCCCAG	ATTTTCCCTG	TATTCCCTGT	ATTCCCTCTA	AAGTAAGCCT	300
AACGAAAATG	TACACGGGTC	TAAAAGGGAC	ATAAGGGACA	TAAGGGAGAT	TTCAATTCGA	
ACACATACAG	GTTGGGCAGA	ATAACAATGT	CTCGAACAAG	GAAAGTGGAC	TCATTACTGC	360
TGTGTATGTC	CAACCCGTCT	TATTGTTACA	GAGCTTGTTT	CTTTCACCTG	AGTAATGACG	
TACTGGCCAT	ACCTGGACTG	GCGCTTCTCT	TATTACCCAA	TGCTTACTGT	GCTTCGTGTG	420
ATGACCGGTA	TGGACCTGAC	CGCGAAGAGA	ATAATGGGTT	ACGAATGACA	CGAAGCACAC	
AGCCTGTGCG	GATCCCCATG	TGCAAATCTA	TGCCATGGAA	CATGACCAAG	ATGCCCAACC	480
TCGGACACGC	CTAGGGGTAC	ACGTTTAGAT	ACGGTACCTT	GTACTGGTTC	TACGGGTTGG	
ATCTCCACCA	CAGCACTCAA	GCCAAATGCCA	TCTTGGAAT	TGAACAGTTT	GAAGGTTTGC	540
TAGAGGTGGT	GTCGTGAGTT	CGGTTACGGT	AGGACCGTTA	ACTTGTCAAA	CTTCCAAACG	
TGACCACTGA	ATGTAGCCAG	GACCTTTTGT	TCTTTCTGTG	TGCCATGTAT	GCCCCCATTT	600
ACTGGTGACT	TACATCGGTC	CTGGAAAACA	AGAAAGACAC	ACGGTACATA	CGGGGGTAAA	
GTACCATCGA	TTTCCAGCAT	GAACCAATTA	AGCCTTGCAA	GTCCGTGTGC	GAAAGGGCCA	660
CATGGTAGCT	AAAGGTCGTA	CTTGGTTAAT	TCGGAACGTT	CAGGCACACG	CTTTCCCGGT	
GGGCCGGCTG	TGAGCCCATT	CTCATAAAGT	ACCGGCACAC	TTGGCCAGAG	AGCCTGGCAT	720
CCCGGCCGAC	ACTCGGGTAA	GAGTATTTC	TGGCCGTGTG	AACCGGTCTC	TCGGACCGTA	
GTGAAGAGCT	GCCCCGTATAT	GACAGAGGAG	TCTGCATCTC	CCCAGAGGCT	ATCGTCACAG	780
CATTTCTCGA	CGGGCATATA	CTGTCTCCTC	AGACGTAGAG	GGGTCTCCGA	TAGCAGTGTC	
TGGAACAAGG	AACAGATTCA	ATGCCAGACT	TCTCCATGGA	TTCAAACAAT	GGAAATTGCG	840
ACCTTGTTCC	TTGTCTAAGT	TACGGTCTGA	AGAGGTACCT	AAGTTTGTTA	CCTTTAACGC	
GAAGCGGCAG	GGAGCACTGT	AAATGCAAGC	CCATGAAGGC	AACCCAAAAG	ACGTATCTCA	900
CTTCGCCGTC	CCTCGTGACA	TTTACGTTTG	GGTACTTCCG	TTGGGTTTTT	TGCATAGAGT	
AGAATAATTA	CAATTATGTA	ATCAGAGCAA	AAGTGAAAGA	GGTGAAAGTG	AAATGCCACG	960
TCTTATTAAT	GTTAATACAT	TAGTCTCGTT	TTCACTTTCT	CCACTTTCAC	TTTACGGTGC	
ACGCAACAGC	AATTGTGGAA	GTAAGGAGA	TTCTCAAGTC	TTCCCTAGTG	AACATTCCCTA	1020
TGCGTTGTG	TTAACACCTT	CATTTCTCT	AAGAGTTCAG	AAGGGATCAC	TTGTAAGGAT	

AAGACACAGT TTCTGTGTCA	GACACTGTAC CTGTGACATG	ACCAACTCAG TGGTTGAGTC	GCTGCTTGTG CGACGAACAC	CCCCCAGCTT GGGGGTCGAA	GTTGCCAATG CAACGGTTAC	1080
AGGAATACAT TCCTTATGTA	AATTATGGGC TTAATACCCG	TATGAAGACA ATACTTCTGT	AAGAGCGTAC TTCTCGCATG	CAGGCTTCTA GTCCGAAGAT	CTAGTGGAAG GATCACCTTC	1140
GATCCTTGGC CTAGGAACCG	CGAAAAATGG GCTTTTTACC	AGAGATCGTC TCTCTAGCAG	TTGCTAAGAA AACGATTCTT	AGTCAAGCGC TCAGTTTCGCG	TGGGATCAAA ACCCTAGTTT	1200
AGCTTCGACG TCGAAGCTGC	TCCCAGGAAA AGGGTCCTTT	AGCAAAGACC TCGTTTCTGG	CCGTGGCTCC GGCACCAGG	AATTCCCAAC TTAAGGGTTG	AAAAACAGCA TTTTTGTCGT	1260
ATTCAGACA TAAGGTCTGT	AGCGCGTAGT TCGCGCATCA	TAGACTAACG ATCTGATTGC	GAAAGGTGTA CTTTCACAT	TGGAACTCT ACCTTTGAGA	ATGGACTTTG TACCTGAAAC	1320
AAACTAAGAT TTTGATTCTA	TTGCATTGTT AACGTAACAA	GGAAGAGCAA CCTTCTCGTT	AAAAGAAATT TTTTCTTTAA	GCACTACAGC CGTGATGTCG	ACGTTATATT TGCAATATAA	1380
CTATTGTTTA GATAACAAAT	CTACAAGAAG GATGTTCTTC	CTGGTTTAGT GACCAAATCA	TGATTGTAGT ACTAACATCA	TCTCCTTTCC AGAGGAAAGG	TTCTTTTTTT AAGAAAAAAA	1440
TTATAACTAT AATATTGATA	ATTTGCACGT TAAACGTGCA	GTTCCCAGGC CAAGGGTCCG	AATTGTTTTA TTAACAAAT	TTCAACTTCC AAGTTGAAGG	AGTGACAGAG TCACTGTCTC	1500
CAGTGACTGA GTCACTGACT	ATGTCTCAGC TACAGAGTCG	CTAAGAAGC GATTTCTTCG	TCAATTCAAT AGTTAAGTAA	TCTGATCAAC AGACTAGTTG	TAATGGTGAC ATTACCACTG	1560
AAGTGTTTGA TTCACAAACT	TACTTGGGGA ATGAACCCCT	AAGTGAAC TA TTCACTTGAT	ATTGCAATGG TAACGTTACC	TAAATCAGAG ATTTAGTCTC	AAAAGTTGAC TTTTCAACTG	1620
CAATGTTGCT GTTACAACGA	TTTCTGTAG AAAGGACATC	ATGAACAAGT TACTTGTTCA	GAGAGATCAC CTCTCTAGTG	ATTTAAATGA TAAATTTACT	TGATCACTTT ACTAGTGAAA	1680
CCATTTAATA GGTAAATTAT	CTTTCAGCAG GAAAGTCGTC	TTTTAGTTAG AAAATCAATC	ATGACATGTA TACTGTACAT	GGATGCACCT CCTACGTGGA	AAATCTAAAT TTTAGATTTA	1740
ATTTTATCAT TAAAATAGTA	AAATGAAGAG TTTACTTCTC	CTGGTTTAGA GACCAAATCT	CTGTATGGTC GACATACCAG	ACTGTTGGGA TGACAACCCCT	AGGTAAATGC TCCATTTACG	1800
CTACTTTGTC GATGAAACAG	AATTCTGTTT TTAAGACAAA	TAAAAATTGC ATTTTAAACG	CTAAATAAAT GATTTATTTA	ATTAAGTCCT TAATTCAGGA	AAATAAAAAA TTTTATTTTT	1860
AAAAAAAAAA TTTTTTTTTT	AAAAA TTTTT					

Fig. 4. (Continuation page 2, SEQ ID NO:4).

MLLLFRAIPM LLLGLMVLQT DCEIAQYYID EEEPPGTVIA VLSQHSIFNT TDIPATNFRL	60
MKQFNNSLIG VRES DGQLSI MERIDREQIC RQSLHCNLAL DVVSFSKGHF KLLNVKVEVR	120
DINDHSPHFP SEIMHVEVSE SSSVGTRIPL EIAIDEDVGS NSIQNFQISN NSHFSIDVLT	180
RADGVKYADL VLMRELDREI QPTYIMELLA MDGGVPSLSG TAVVNIRVLD FNDNSPVFER	240
STIAVDLVED APLGYLLEL HATDDDEGVN GEIVYGFSTL ASQEVROLFK INSRTGSVTL	300
EGQVDFETKQ TYEFEVQAQD LGPNPLTATC KVTVHILDVN DNTPAITITP LTTVNAGVAY	360
IPETATKENF IALISTTDRA SGSNGQVRCT LYGHEHFKLQ QAYEDSYMIV TTSTLDRENI	420
AAYSLTVVAE DLGFPSLGTK KYITVKVSDE NDNAPVFSKP QYEASILENN APGSYITTVI	480
ARDSDSQNG KVNRYLVDK VMGQSLTTFV SLDADSGVLR AVRSLDYEKL KQLDFEIEAA	540
DNGIPQLSTR VQLNLRIVDQ NDNCPVITNP LLNNGSGEVL LPISAPQNYL VFQLKAEDSD	600
EGHNSQLFYT ILRDPSRLFA INKESGEVFL KKQLNSDHSE DLSIVVAVYD LGRPSLSTNA	660
TVKFILTDSF PSNVEVILQ PSAEEQHQID MSIIFIAVLA GGCALLLLAI FFVACTCKKK	720
AGEFKQVPEQ HGTCNEERLL STPSPQSVSS SLSQSESCQL SINTESENCV VSSNQEQHQQ	780
TGIKHSISVP SYHTSGWHLN NCAMSSGHG HMGHISTKVQ WAKEIVTSMT VTLILVENQK	840
RRALSSQCRH KPVLTNTQMNQ QGSDMPITIS ATESTRVQKM GTAHCMNKRA IDCLTL	

Figure 5. Deduced amino acid sequence of the *Xenopus* PAPC (paraxial protocadherin) protein. It encodes a member of the cadherin family of transmembrane proteins that has dorsalizing activity when constructs are injected into *Xenopus* embryos. SEQ ID NO:5.

Figure 6. Nucleotide sequence of the full-length PAPC cDNA derived from the *Xenopus* organizer. The sense strand of the DNA is shown in the top line (in the 5' to 3' direction), and the bottom line shows the antisense strand (opposite orientation). SEQ ID NO:6.

GAATTC	CCAG	AGATGA	ACTC	CTTGAG	ATTG	TTTTAA	TGA	CTGCAG	GTCT	GGAAGG	ATTC	60
CTTAAG	GGTC	TCTACT	TGAG	GAAC	TCTAAC	AAAATT	TACT	GACG	TCCAGA	CCTT	CTTAAG	
ACATTG	CCAC	ACTG	TTTCTA	GGCATG	AAAA	AACTGC	AAAGT	TTCAAC	TTTG	TTTTGG	TGC	120
TGTAAC	GGTG	TGACAA	AGAT	CCGTACT	TTT	TTGACG	TTCA	AAGTTG	AAAC	AAAAAC	CACG	
AACTTT	GATT	CTTCAAG	ATG	CTGCTT	CTCT	TCAGAG	CCCAT	TCCAAT	GC	CTGT	TGGGAC	180
TTGAA	ACTAA	GAAGTT	CTAC	GACGA	AGAG	AGTCTC	GGTA	AGGTTA	CGAC	GACAAC	CCCTG	
TGATGG	TTTT	ACAAAC	AGAC	TGTGAA	AATTG	CCCAGT	ACTA	CATAGAT	GAA	GAAGA	ACCCC	240
ACTAC	CAAAA	TGTTT	GTCTG	ACACTT	TAAAC	GGGTCAT	GAT	GTATCT	ACTT	CTTCTT	TGGGG	
CTGGCA	CTGT	AATTGC	AGTG	TTGTC	CACAAC	ACTCC	ATATT	TAACACT	TACA	GATATA	CCCTG	300
GACCGT	GACA	TTAACG	TCAC	AACAGT	GTG	TGAGGT	TATAA	ATTGTG	ATGT	CTATAT	TGGAC	
CAACCA	AATTT	CCGTCT	AATG	AAGCA	AATTTA	ATAATT	CCCT	TATCGG	AGTC	CGTGAG	AGTG	360
GTTGGT	TAAA	GGCAG	ATTAC	TTCGT	TAAAT	TATTA	AGGGA	ATAGCC	TCAG	GCACTC	TCAC	
ATGGGC	AGCT	GAGCAT	CATG	GAGAGG	ATTG	ACCGGG	AGCA	AATCTG	CAGG	CAGTCC	CTTC	420
TACCCG	TCGA	CTCGTA	GTAC	CTCTC	CTAAC	TGGCC	CTCGT	TTAGAC	GTCC	GTCAGG	GAAG	
ACTGCA	ACCT	GGCTTT	TGGAT	TGGTC	AGCT	TTTCCA	AAAG	ACACTT	CAAG	CTTCTG	AACG	480
TGACGT	TGGA	CCGAA	ACCTA	CACCAG	TCGA	AAAGG	TTTCC	TGTGA	AGTTC	GAAGAC	TTGC	
TGAAAG	TGGA	GGTGAG	AGAC	ATTAAT	AGCC	ATAGCC	CTCA	CTTTCC	CAGT	GAAATA	ATGC	540
ACTTTC	ACCT	CCACTC	TCTG	TAATT	ACTG	TATCGG	GAGT	GAAAGG	GTCA	CTTTAT	TACG	
ATGTGG	AGGT	GTCTGA	AAAGT	TCCTCT	GTGG	GCACC	AGGAT	TCCTTT	AGAA	ATTGCA	ATAG	600
TACAC	CTCA	CAGACT	TTTCA	AGGAG	ACACC	CGTGG	TCCTA	AGGAA	ATCTT	TAACGT	TATC	
ATGAAG	ATGT	TGGGT	CCAAC	TCCAT	CCAGA	ACTTT	CAGAT	CTCAA	ATAAT	AGCCAC	TTCA	660
TACTT	CTACA	ACCCAG	GTTG	AGGTAG	GTCT	TGAAAG	TCTA	GAGTTT	TATTA	TCGGT	GAAAT	
GCATTG	ATGT	GCTAAC	CCAGA	GCAGAT	TGGGG	TGAAAT	ATATGC	AGATTT	AGTC	TTAATG	AGAG	720
CGTAA	CTACA	CGATTG	GTCT	CGTCT	ATCCCC	ACTTT	TATACG	TCTAA	ATCAG	AATTAC	TCTC	
AACTGG	ACAG	GGAAAT	CCAG	CCAAC	ATACA	TAATGG	AGCT	ACTAG	CAATG	GATGGG	GGTG	780
TTGAC	CTGTC	CCTTT	AGGTC	GGTTG	TATGT	ATTAC	CTCGA	TGATC	GTTAC	CTACCC	CCAC	
TACCAT	CACT	ATCTGG	TACT	GCAGT	TGGTTA	ACATCC	GAGT	CCTGG	ACTTT	AATGAT	AACA	840
ATGGT	AGTGA	TAGACC	ATGA	CGTCA	CCAAT	TGTAGG	CTCA	GGAC	CTGAA	TTACT	ATTGT	
GCCCAG	TGTT	TGAGAG	AAGC	ACCATT	TGCTG	TGGAC	CTAGT	AGAGG	ATGCT	CCTCT	TGGGAT	900
CGGGT	CACAA	ACTCT	CTTCG	TGGTA	ACGAC	ACCTGG	ATCA	TCTC	CTACGA	GGAGAC	CCCTA	
ACCTTT	TGTT	GGAGT	TACAT	GCTACT	GACG	ATGAT	GAAGG	AGTGA	ATGGA	GAAATT	TGTTT	960
TGGAAA	ACAA	CCTCA	ATGTA	CGATG	ACTGC	TACTAC	TTCC	TCAC	TTACCT	CTTTA	ACAAA	
ATGGAT	TCAG	CACTTT	TGCA	TCTCA	AGAGG	TACG	TCAGT	ATTTAA	AAATT	AACTCC	AGAA	1020
TACCTA	AGTC	GTGAA	ACCGT	AGAGT	TCTCC	ATGC	AGTCGA	TAAAT	TTTTAA	TTGAGG	TCTT	

CTGGCAGTGT GACCGTCACA	TACTCTTGAA ATGAGAACTT	GGCCAAGTTG CCGGTTCAAC	ATTTTGAGAC TAAAACTCTG	CAAGCAGACT GTTCTGCTGA	TACGAATTTG ATGCTTAAAC	1080
AGGTACAAGC TCCATGTTCT	CCAAGATTTG GGTTCATAAC	GGCCCCAACC CCGGGGTTGG	CACTGACTGC GTGACTGACG	TACTTGTAAA ATGAACATTT	GTAAGTGTTC CATTGACAAG	1140
ATATACTTGA TATATGAACT	TGTAAATGAT ACATTTACTA	AATACCCAG TTATGGGGTC	CCATCACTAT GGTAGTGATA	TACCCCTCTG ATGGGGAGAC	ACTACTGTAA TGATGACATT	1200
ATGCAGGAGT TACGTCTCTA	TGCCTATATT ACGGATATAA	CCAGAAACAG GGTCTTTGTC	CCACAAAGGA GGTGTTCCTT	GAACTTTATA CTTGAAATAT	GCTCTGATCA CGAGACTAGT	1260
GCACTACTGA CGTGATGACT	CAGAGCCTCT GTCTCGGAGA	GGATCTAATG CCTAGATTAC	GACAAGTTCT CTGTTCAGAC	CTGTACTCTT GACATGAGAA	TATGGACATG ATACCTGTAC	1320
AGCACTTTAA TCGTGAAATT	ACTACAGCAA TGATGTCGTT	GCTTATGAGG CGAATACTCC	ACAGTTACAT TGTCAATGTA	GATAGTTACC CTATCAATGG	ACCTCTACTT TGGAGATGAA	1380
TAGACAGGGA ATCTGTCCCT	AAACATAGCA TTTGTATCGT	GCGTACTCTT CGCATGAGAA	TGACAGTAGT ACTGTCATCA	TGCAGAAGAC ACGTCTTCTG	CTTGGCTTCC GAACCGAAGG	1440
CCTCATTGAA GGAGTAACTT	GACCAAAAAG CTGGTTTTTC	TACTACACAG ATGATGTGTC	TCAAGGTTAG AGTTCCAATC	TGATGAGAAT ACTACTCTTA	GACAATGCAC CTGTTACGTG	1500
CTGTATTTTC GACATAAAAG	TAAACCCAG ATTTGGGGTC	TATGAAGCTT ATACTTCGAA	CTATTCTGGA GATAAGACCT	AAATAATGCT TTTATTACGA	CCAGGCTCTT GGTCCGAGAA	1560
ATATAACTAC TATATTGATG	AGTGATAGCC TCACTATCGG	AGAGACTCTG TCTCTGAGAC	ATAGTGATCA TATCACTAGT	AAATGGCAAA TTTACCGTTT	GTAATTAACA CATTTAATGT	1620
GACTTGTGGA CTGAACACCT	TGCAAAAGTG ACGTTTTTCAC	ATGGGCCAGT TACCCGGTCA	CACTAACAAC GTGATTGTTG	ATTTGTTTCT TAAACAAAGA	CTTGATGCGG GAACTACGCC	1680
ACTCTGGAGT TGAGACCTCA	ATTGAGAGCT TAACTCTCGA	GTTAGGTCTT CAATCCAGAA	TAGACTATGA ATCTGATACT	AAACTTTAAA TTTTGAATTT	CAACTGGATT GTTGACCTAA	1740
TTGAAATTGA AACTTTAACT	AGCTGCAGAC TCGACGTCTG	AATGGGATCC TTACCCTAGG	CTCAACTCTC GAGTTGAGAG	CACTCGCGTT GTGAGCGCAA	CAACTAAATC GTTGATTTAG	1800
TCAGAAATAGT AGTCTTATCA	TGATCAAAAT ACTAGTTTTA	GATAATTGCC CTATTAACGG	CTGTGATAAC GACACTATTG	TAATCCTCTT ATTAGGAGAA	CTTAATAATG GAATTATTAC	1860
GCTCGGGTGA CGAGCCCACT	AGTTCCTGCT TCAAGACGAA	CCCATCAGCG GGGTAGTCGC	CTCCTCAAAA GAGGAGTTTT	CTATTTAGTT GATAAATCAA	TTCCAGCTCA AAGGTCGAGT	1920
AAGCCGAGGA TTCGGCTCCT	TTCAGATGAA AAGTCTACTT	GGGCACAAC CCCGTGTTGA	CCCAGCTGTT GGGTCGACAA	CTATACCATA GATATGGTAT	CTGAGAGATC GACTCTCTAG	1980
CAAGCAGATT GTTCTGCTAA	GTTTGCCATT CAAACGGTAA	AACAAAGAAA TTGTTTCTTT	GTGGTGAAAT CACCACCTTA	GTTCTGAAA CAAGGACTTT	AAACAATTAA TTTGTTAATT	2040
ACTCTGACCA TGAGACTGGT	TTCAGAGGAC AAGTCTCCTG	TTGAGCATAG AACTCGTATC	TAGTTGCAGT ATCAACGTCA	GTATGACTTG CATACTGAAC	GGAAGACCTT CCTTCTGGAA	2100
CATTATCCAC GTAATAGGTG	CAATGCTACA GTTACGATGT	GTTAAATTCA CAATTTAAGT	TCCTCACCGA AGGAGTGGCT	CTCTTTTCTT GAGAAAAGGA	TCTAACGTTG AGATTGCAAC	2160

Fig. 6. (Continuation page 2, SEQ ID NO:6).

AAGTCGTTAT TTTGCAACCA TCTGCAGAAG AGCAGCACCA GATCGATATG TCCATTATAT	2220
TTCAGCAATA AAACGTTGGT AGACGTCTTC TCGTCGTGGT CTAGCTATAC AGGTAATATA	
TCATTGCAGT GCTGGCTGGT GGTGTGTGCTT TGCTACTTTT GGCCATCTTT TTTGTGGCCT	2280
AGTAACGTCA CGACCGACCA CCAACACGAA ACGATGAAAA CCGGTAGAAA AAACACCGGA	
GTACTTGTA AAGAAAGCT GGTGAATTTA AGCAGGTACC TGAACAACAC GGAACATGCA	2340
CATGAACATT TTTCTTTTGA CCACTTAAAT TCGTCCATGG ACTTGTGTG CCTGTACGT	
ATGAAGAACG CCTGTTAAGC ACCCATCTC CCCAGTCGGT CTCTTCTTCT TTGTCTCAGT	2400
TACTTCTTGC GGACAATTCG TGGGGTAGAG GGGTCAGCCA GAGAAGAAGA AACAGAGTCA	
CTGAGTCATG CCAACTCTCC ATCAATACTG AATCTGAGAA TTGCAGCGTG TCCTCTAACC	2460
GACTCAGTAC GGTGAGAGG TAGTTATGAC TTAGACTCTT AACGTCGCAC AGGAGATTGG	
AAGAGCAGCA TCAGCAAACA GGCATAAAGC ACTCCATCTC TGTACCATCT TATCACACAT	2520
TTCTCGTCGT AGTCGTTTGT CCGTATTTCT TGAGGTAGAG ACATGGTAGA ATAGTGTGTA	
CTGGTTGGCA CCTGGACAAT TGTGCAATGA GCATAAGTGG ACATTCTCAC ATGGGGCACA	2580
GACCAACCGT GGACCTGTTA ACACGTTACT CGTATTCACC TGTAAGAGTG TACCCCGTGT	
TTAGTACAAA GGTACAGTGG GCAAAGGAGA TAGTGACTTC AATGACAGTG ACTCTGATAC	2640
AATCATGTTT CCATGTCACC CGTTTCCTCT ATCACTGAAG TTACTGTCAC TGAGACTATG	
TAGTGGAGAA TCAGAAAAGA AGAGCATTGA GCAGCCAATG CAGGCACAAG CCAGTGCTCA	2700
ATCACCTCTT AGTCTTTTCT TCTCGTAACT CGTCGGTTAC GTCCGTGTTC GGTACAGAGT	
ATACACAGAT GAATCAGCAG GGTTCGACA TGCCGATAAC TATTTTCAGCC ACCGAATCAA	2760
TATGTGTCTA CTTAGTCGTC CCAAGGCTGT ACGGCTATTG ATAAAGTCGG TGGCTTAGTT	
CAAGGGTCCA GAAATGGA ACTGCACATT GCAATATGAA AAGGGCTATA GACTGTCTTA	2820
GTTCCAGGT CTTTACCCT TGACGTGTAA CGTTATACTT TTCCCGATAT CTGACAGAAT	
CTCTGTAGCT CCTGTATATT ACAATACCTA CCATGCAAGA ATGCCTAACC TGCACATACC	2880
GAGACATCGA GGACATATAA TGTATGGAT GGTACGTTCT TACGGATTGG ACGTGTATGG	
GAACCATACC CTTAGAGACC CTTATTACCA TATCAATAAT CCTGTTGCTA ATCGGATGCA	2940
CTTGGTATGG GAATCTCTGG GAATAATGGT ATAGTTATTA GGACAACGAT TAGCCTACGT	
GGCGGAATAT GAAAGAGATT TAGTCAACAG AAGTGCAACG TTATCTCCGC AGAGATCGTC	3000
CCGCCTTATA CTTTCTCTAA ATCAGTTGTC TTCACGTTGC AATAGAGGCG TCTCTAGCAG	
TAGCAGATAC CAAGAATTCA ATTACAGTCC GCAGATATCA AGACAGCTTC ATCCTTCAGA	3060
ATCGTCTATG GTTCTTAAGT TAATGTCAGG CGTCTATAGT TCTGTGCAAG TAGGAAGTCT	
AATTGCTACA ACCTTTTAAT CATTAGGCAT GCAAGTGAGA ATGCACAAAG GCAAGTGCTT	3120
TTAACGATGT TGGAAAATTA GTAATCCGTA CGTTCACCTT TACGTGTTTC CGTTCACGAA	
TAGCATGAAA GCTAAATATA TGGAGTCTCC CCTTCCCTC TGATGGATGG GGGGAGACAC	3180
ATCGTACTTT CGATTTATAT ACCTCAGAGG GGAAAGGGAG ACTACCTACC CCCCTCTGTG	
AGGACAGTGC ATAAATATAC AGCTGCTTTC TATTGTCATT TCACTTGGGA ATTTTTTGT	3240
TCCTGTCACG TATTTATATG TCGACGAAAG ATAAACGTAA AGTGAACCCT TAAAAACAA	
TTTTTTACAT ATTTATTTTT CCTGAATTGA ATGTGACATT GTCCTGTCAC CTAAGTAGCA	3300
AAAAAATGTA TAAATAAAAA GGACTTAACT TACACTGTAA CAGGACAGTG GATTGATCGT	

Fig. 6. (Continuation page 3, SEQ ID NO:6).

MVCCGPGRML LGWAGLLVLA ALCLLQVPGA QAAACEPVRI PLCKSLPWNM TKMPNHLHHS 60
TQANAILAME QFEGLLGTHC SPDLLFFLCA MYAPICTIDF QHEPIKPCKS VCERARQGCE 120
PILIKYRHSW PESLACDELP VYDRGVCISP EAIVTADGAD FPMDSSTGHC RGASSERCKC 180
KPVRATQKTY FRNNYNYVIR AKVKEVKMKC HDVTAVVEVK EILKASLVNI PRDTVNLVTT 240
SGCLCPPLTV NEEYVIMGYE DEERSRLLLV EGSIAEKWKD RLGKKVKRWD MKLRHLGLGK 300
TDASDSTQNQ KSGRNSNPRP ARS.

Figure 7. Deduced amino acid sequence of mouse FRZB-1 protein. SEQ ID NO:7.

Figure 8. Nucleotide sequence of the full-length mouse FRZB-1 cDNA. SEQ ID NO:8.

AAGCCTGGGA	CCATGGTCTG	CTGCGGCCCCG	GGACGGATGC	TGCTAGGATG	GGCCGGGTTG	60
TTCGGACCCCT	GGTACCAGAC	GACGCCGGGC	CCTGCCTACG	ACGATCCTAC	CCGGCCCAAC	
CTAGTCCTGG	CTGCTCTCTG	CCTGCTCCAG	GTGCCCCGAG	CTCAGGCTGC	AGCCTGTGAG	120
GATCAGGACC	GACGAGAGAC	GGACGAGGTC	CACGGGCCTC	GAGTCCGACG	TCGGACACTC	
CCTGTCCGCA	TCCCCGCTGTG	CAAGTCCCTT	CCCTGGAACA	TGACCAAGAT	GCCCAACCAC	180
GGACAGGCGT	AGGGCGACAC	GTTCAGGGAA	GGGACCTTGT	ACTGGTTCTA	CGGGTTGGTG	
CTGCACCACA	GCACCCAGGC	TAACGCCATC	CTGGCCATGG	AACAGTTCGA	AGGGCTGCTG	240
GACGTGGTGT	CGTGGGTCCG	ATTGCGGTAG	GACCGGTACC	TTGTCAAGCT	TCCCACGAC	
GGCACCCACT	GCAGCCCGGA	TCTTCTCTTC	TTCCTCTGTG	CAATGTACGC	ACCCATTTGC	300
CCGTGGGTGA	CGTCGGGCCT	AGAAGAGAAG	AAGGAGACAC	GTTACATGCG	TGGGTAAACG	
ACCATCGACT	TCCAGCACGA	GCCCATCAAG	CCCTGCAAGT	CTGTGTGTGA	GCGCGCCCGA	360
TGGTAGCTGA	AGGTCGTGCT	CGGGTAGTTC	GGGACGTTCA	GACACACACT	CGCGCGGGCT	
CAGGGCTGCG	AGCCCATTCT	CATCAAGTAC	CGCCACTCGT	GGCCGGAAAAG	CTTGGCCTGC	420
GTCCCGACGC	TCGGGTAAGA	GTAGTTTCATG	GCGGTGAGCA	CCGGCCTTTC	GAACCGGACG	
GACGAGCTGC	CGGTGTACGA	CCGCGGCGTG	TGCATCTCTC	CTGAGGCCAT	CGTCACCGCG	480
CTGCTCGACG	GCCACATGCT	GGCGCCGCAC	ACGTAGAGAG	GACTCCGGTA	GCAGTGGCGC	
GACGGAGCGG	ATTTTCCTAT	GGATTCAAGT	ACTGGACACT	GCAGAGGGGC	AAGCAGCGAA	540
CTGCCTCGCC	TAAAAGGATA	CCTAAGTTCA	TGACCTGTGA	CGTCTCCCCG	TTCGTCGCTT	
CGTTGCAAAT	GTAAGCCTGT	CAGAGCTACA	CAGAAGACCT	ATTTCCGGAA	CAATTACAAC	600
GCAACGTTTA	CATTTCGGACA	GTCTCGATGT	GTCTTCTGGA	TAAAGGCCTT	GTTAATGTTG	
TATGTCATCC	GGGCTAAAGT	TAAAGAGGTA	AAGATGAAAT	GTCATGATGT	GACCGCCGTT	660
ATACAGTAGG	CCCGATTTC	ATTTCTCCAT	TTCTACTTTA	CAGTACTACA	CTGGCGGCAA	
GTGGAAGTGA	AGGAAATTCT	AAAGGCATCA	CTGGTAAACA	TTCCAAGGGA	CACCGTCAAT	720
CACCTTCACT	TCCTTTAAGA	TTTCCGTAGT	GACCATTGT	AAGGTTCCCT	GTGGCAGTTA	
CTTTATACCA	CCTCTGGCTG	CCTCTGTCCT	CCACTTACTG	TCAATGAGGA	ATATGTCATC	780
GAAATATGGT	GGAGACCGAC	GGAGACAGGA	GGTGAATGAC	AGTTACTCCT	TATACAGTAG	
ATGGGCTATG	AAGACGAGGA	ACGTTCCAGG	TTACTCTTGG	TAGAAGGCTC	TATAGCTGAG	840
TACCCGATAC	TTCTGCTCCT	TGCAAGGTCC	AATGAGAACC	ATCTTCCGAG	ATATCGACTC	
AAGTGGAAGG	ATCGGCTTGG	TAAGAAAGTC	AAGCGCTGGG	ATATGAAACT	CCGACACCTT	900
TTCACCTTCC	TAGCCGAACC	ATTCTTTTCAG	TTCGCGACCC	TATACTTTGA	GGCTGTGGAA	
GGACTGGGTA	AAACTGATGC	TAGCGATTCC	ACTCAGAATC	AGAAGTCTGG	CAGGAACTCT	960
CCTGACCCAT	TTTGACTACG	ATCGCTAAGG	TGAGTCTTAG	TCTTCAGACC	GTCTTGTAGA	

AATCCCCGGC	CAGCACGCAG	CTAAATCCTG	AAATGTAAAA	GGCCACACCC	ACGGACTCCC	1020
TTAGGGGCGG	GTCGTGCGTC	GATTTAGGAC	TTTACATTTT	CCGGTGTGGG	TGCCTGAGGG	
TTCTAAGACT	GGCGCTGGTG	GAATAACAAA	GGAAAACCGC	ACAGTTGTGC	TCGTGACCGA	1080
AAGATTCTGA	CCGCGACCAC	CTGATTGTTT	CCTTTTGGCG	TGTCAACACG	AGCACTGGCT	
TTGTTTACCG	CAGACACCGC	GTGGCTACCG	AAGTTACTTC	CGGTCCCCTT	TCTCCTGCTT	1140
AACAAATGGC	GTCTGTGGCG	CACCGATGGC	TTCAATGAAG	GCCAGGGGAA	AGAGGACGAA	
CTTAATGGCG	TGGGGTTAGA	TCCTTTAATA	TGTTATATAT	TCTGTTTCAT	CAATCACGTG	1200
GAATTACCGC	ACCCCAATCT	AGGAAATTAT	ACAATATATA	AGACAAAGTA	GTTAGTGCAC	
GGGACTGTTC	TTTTGCAACC	AGAATAGTAA	ATTAAATATG	TTGATGCTAA	GGTTTCTGTA	1260
CCCTGACAAG	AAAACGTTGG	TCTTATCATT	TAATTTATAC	AACTACGATT	CCAAAGACAT	
CTGGACTCCC	TGGGTTTAAT	TTGGTGTTCT	GTACCCTGAT	TGAGAATGCA	ATGTTTCATG	1320
GACCTGAGGG	ACCCAAATTA	AACCACAAGA	CATGGGACTA	ACTCTTACGT	TACAAAGTAC	
TAAAGAGAGA	ATCCTGGTCA	TATCTCAAGA	ACTAGATATT	GCTGTAAGAC	AGCCTCTGCT	1380
ATTTCTCTCT	TAGGACCAGT	ATAGAGTTCT	TGATCTATAA	CGACATTCTG	TCGGAGACGA	
GCTGCGCTTA	TAGTCTTG TG	TTTGTATGCC	TTTGTCCATT	TCCCTCATGC	TGTGAAAGTT	1440
CGACGCGAAT	ATCAGAACAC	AAACATACGG	AAACAGGTAA	AGGGAGTACG	ACACTTTCAA	
ATACATGTTT	ATAAAGGTAG	AACGGCATT	TGAAATCAGA	CACTGCACAA	GCAGAGTAGC	1500
TATGTACAAA	TATTTCCATC	TTGCCGTAAA	ACTTTAGTCT	GTGACGTGTT	CGTCTCATCG	
CCAACACCAG	GAAGCATTTA	TGAGGAAACG	CCACACAGCA	TGACTTATTT	TCAAGATTGG	1560
GGTTGTGGTC	CTTCGTAAAT	ACTCCTTTGC	GGTGTGTCGT	ACTGAATAAA	AGTTCTAACC	
CAGGCAGCAA	AATAAATAGT	GTTGGGAGCC	AAGAAAAGAA	TATTTTGCCT	GGTTAAGGGG	1620
GTCCGTCGTT	TTATTTATCA	CAACCCTCGG	TTCTTTTCTT	ATAAACGGA	CCAATTCCCC	
CACACTGGAA	TCAGTAGCCC	TTGAGCCATT	AACAGCAGTG	TTCTTCTGGC	AAGTTTTTGA	1680
GTGTGACCTT	AGTCATCGGG	AACTCGGTAA	TTGTCGTCAC	AAGAAGACCG	TTCAAAAAC	
TTTGTTTATA	AATGTATTCA	CGAGCATTAG	AGATGAACTT	ATAACTAGAC	ATCTGTTGTT	1740
AAACAAGTAT	TTACATAAGT	GCTCGTAATC	TCTACTTGAA	TATTGATCTG	TAGACAACAA	
ATCTCTATAG	CTCTGCTTCC	TTCTAAATCA	AACCCATTGT	TGGATGCTCC	CTCTCCATTC	1800
TAGAGATATC	GAGACGAAGG	AAGATTTAGT	TTGGGTAACA	ACCTACGAGG	GAGAGGTAAG	

09903345-074101

Figure 10. Nucleotide sequence of the full-length human FRZB-1 cDNA. SEQ ID NO:10.
This sequence was assembled from public ESTs from the Genbank database
(accession numbers: H18848, R63748, W38677, W44760, H38379 and N71244).

GGCGGAGCGG GCCTTTTGGC GTCCACTGCG CGGCTGCACC CTGCCCCATC TGCCGGGATC	60
CCGCCTCGCC CGGAAAACCG CAGGTGACGC GCCGACGTGG GACGGGGTAG ACGGCCCTAG	
ATGGTCTGCG GCAGCCCGGG AGGGATGCTG CTGCTGCGGG CCGGGCTGCT TGCCCTGGCT	120
TACCAGACGC CGTCGGGCCC TCCCTACGAC GACGACGCCC GGCCCGACGA ACGGGACCGA	
GCTCTCTGCC TGCTCCGGGT GCCCGGGGCT CGGGCTGCAG CCTGTGAGCC CGTCCGCATC	180
CGAGAGACGG ACGAGGCCCA CGGGCCCCGA GCCCGACGTC GGACACTCGG GCAGGCGTAG	
CCCCTGTGCA AGTCCCTGCC CTGGAACATG ACTAAGATGC CCAACCACCT GCACCACAGC	240
GGGGACACGT TCAGGGACGG GACCTTGTA TATTCTACG GGTGGGTGGA CGTGGTGTCTG	
ACTCAGGCCA ACGCCATCCT GGCCATCGAG CAGTTCGAAG GTCTGCTGGG CACCCACTGC	300
TGAGTCCGGT TCGGGTAGGA CCGGTAGCTC GTCAAGCTTC CAGACGACCC GTGGGTGACG	
AGCCCCGATC TGCTCTTCTT CCTCTGTGCC ATGTACGCGC CCATCTGCAC CATTGACTTC	360
TCGGGGCTAG ACGAGAAGAA GGAGACACGG TACATGCGCG GGTAGACGTG GTAAGTGAAG	
CAGCACGAGC CCATCAAGCC CTGTAAAGTCT GTGTGCGAGC GGGCCCCGCA GGGCTGTGAG	420
GTCGTGCTCG GGTAGTTTCGG GACATTTCAG CACACGCTCG CCCGGGGCGT CCCGACACTC	
CCCATACTCA TCAAGTACCG CCACTCGTGG CCGGAGAACC TGGCCTGCGA GGAGCTGCCA	480
GGGTATGAGT AGTTCATGGC GGTGAGCACC GGCCTCTTGG ACCGGACGCT CCTCGACGGT	
GTGTACGACA GGGGCGTGTG CATCTCTCCC GAGGCCATCG TTAGTGCAGG CGGAGCTGAT	540
CACATGCTGT CCCCACACAC GTAGAGAGGG CTCCGGTAGC AATGACGCCT GCCTCGACTA	
TTTCCTATGG ATTCTAGTAA CGGAAACTGT AGAGGGGCAA GCAGTGAACG CTGTAAATGT	600
AAAGGATACC TAAGATCATT GCCTTTGACA TCTCCCCGTT CGTCACTTGC GACATTTACA	
AAGCCTATTA GAGCTACACA GAAGACCTAT TTCCGGAACA ATTACAATA TGTCATTCCG	660
TTCGGATAAT CTCGATGTGT CTTCTGGATA AAGGCCTTGT TAATGTTGAT ACAGTAAGCC	
GCTAAAGTTA AAGAGATAAA GACTAAGTGC CATGATGTGA CTGCAGTAGT GGAGGTGAAG	720
CGATTTCAAT TTCTCTATTT CTGATTCACG GTACTACACT GACGTCATCA CCTCCACTTC	
GAGATTCTAA AGTCCTCTCT GGTAAACATT CCACGGGACA CTGTCAACCT CTATACCAGC	780
CTCTAAGATT TCAGGAGAGA CCATTTGTAA GGTGCCCTGT GACAGTTGGA GATATGGTCG	
TCTGGCTGCC TCTGCCCTCC ACTTAATGTT AATGAGGAAT ATATCATCAT GGGCTATGAA	840
AGACCGACGG AGACGGGAGG TGAATTACAA TTAATCCTTA TATAGTAGTA CCCGATACTT	

GATGAGGAAC	GTTCCAGATT	ACTCTTGGTG	GAAGGCTCTA	TAGCTGAGAA	GTGGAAGGAT	900
CTACTCCTTG	CAAGGTCTAA	TGAGAACCAC	CTTCCGAGAT	ATCGACTCTT	CACCTTCCTA	
CGACTCGGTA	AAAAAGTTAA	GCGCTGGGAT	ATGAAGCTTC	GTCATCTTGG	ACTCAGTAAA	960
GCTGAGCCAT	TTTTTCAATT	CGCGACCCTA	TACTTCGAAG	CAGTAGAACC	TGAGTCATTT	
AGTGATTCTA	GCAATAGTGA	TTCCACTCAG	AGTCAGAAGT	CTGGCAGGAA	CTCGAACCCC	1020
TCACTAAGAT	CGTTATCACT	AAGGTGAGTC	TCAGTCTTCA	GACCGTCCTT	GAGCTTGGGG	
CGGCAAGCAC	GCAACTAAAT	CCCGAAATAC	AAAAAGTAAC	ACAGTGGACT	TCCTATTAAG	1080
GCCGTTCGTG	CGTTGATTTA	GGGCTTTATG	TTTTTCATTG	TGTCACCTGA	AGGATAATTC	
ACTTACTTGC	ATTGCTGGAC	TAGCAAAGGA	AAATTGCACT	ATTGCACATC	ATATTCTATT	1140
TGAATGAACG	TAACGACCTG	ATCGTTTCCT	TTTAACGTGA	TAACGTGTAG	TATAAGATAA	
GTTTACTATA	AAAATCATGT	GATAACTGAT	TATTACTTCT	GTTTCTCTTT	TGGTTTCTGC	1200
CAAATGATAT	TTTTAGTACA	CTATTGACTA	ATAATGAAGA	CAAAGAGAAA	ACCAAAGACG	
TTCTCTCTTC	TCTCAACCCC	TTTGTAATGG	TTTGGGGGCA	GACTCTTAAG	TATATTGTGA	1260
AAGAGAGAAG	AGAGTTGGGG	AAACATTACC	AAACCCCCGT	CTGAGAATTC	ATATAACACT	
GTTTTCTATT	TCACTAATCA	TGAGAAAAAC	TGTTCTTTTG	CAATAATAAT	AAATTAAACA	1320
CAAAAGATAA	AGTGATTAGT	ACTCTTTTTG	ACAAGAAAAAC	GTTATTATTA	TTTAATTTGT	
TGCTGTTACC	AGAGCCTCTT	TGCTGAGTCT	CCAGATGTTA	ATTTACTTTC	TGCACCCCAA	1380
ACGACAATGG	TCTCGGAGAA	ACGACTCAGA	GGTCTACAAT	TAAATGAAAAG	ACGTGGGGTT	
TTGGGAATGC	AATATTGGAT	GAAAAGAGAG	GTTTCTGGTA	TTACACAGAAA	GCTAGATATG	1440
AACCCTTACG	TTATAACCTA	CTTTTCTCTC	CAAAGACCAT	AAGTGTCTTT	CGATCTATAC	
CCTTAAACA	TACTCTGCCG	ATCTAATTAC	AGCCTTATTT	TTGTATGCCT	TTTGGGCATT	1500
GGAATTTTGT	ATGAGACGGC	TAGATTAATG	TCGGAATAAA	AACATACGGA	AAACCCGTAA	
CTCCTCATGC	TTAGAAAGTT	CCAAATGTTT	ATAAAGGTAA	AATGGCAGTT	TGAAGTCAAA	1560
GAGGAGTACG	AATCTTTCAA	GGTTTACAAA	TATTTCCATT	TTACCGTCAA	ACTTCAGTTT	
TGTCACATAG	GCAAAGCAAT	CAAGCACCAG	GAAGTGTTTA	TGAGGAAAACA	ACACCCAAGA	1620
ACAGTGATATC	CGTTTCGTGA	GTTTCGTGGTC	CTTCACAAAT	ACTCCTTTGT	TGTGGGTTCCT	
TGAATTATTT	TTGAGACTGT	CAGGAAGTAA	AATAAATAGG	AGCTTAAGAA	AGAACATTTT	1680
ACTTAATAAA	AACTCTGACA	GTCCTTCATT	TTATTTATCC	TCGAATTCTT	TCTTGTAAAA	
GCCTGATTGA	GAAGCACAAAC	TGAAACCAGT	AGCCGCTGGG	GTGTTAATGG	TAGCATTTCTT	1740
CGGACTAACT	CTTCGTGTTG	ACTTTGGTCA	TCGGCGACCC	CACAATTACC	ATCGTAAGAA	
CTTTTGGCAA	TACATTTGAT	TTGTTTCATGA	ATATATTAAT	CAGCATTAGA	GAAATGAATT	1800
GAAAACCGTT	ATGTAAACTA	AACAAGTACT	TATATAATTA	GTCGTAATCT	CTTTACTTAA	
ATAACTAGAC	ATCTGCTGTT	ATCACCATAG	TTTTGTTTAA	TTTGCTTCCT	TTTAAATAAA	1860
TATTGATCTG	TAGACGACAA	TAGTGGTATC	AAAACAAATT	AAACGAAGGA	AAATTTATTT	
CCCATTTGGTG	AAAGTCAAAA	AAAAAAAAAA	AAA			
GGGTAACCAC	TTTCAGTTTT	TTTTTTTTTT	TTT			